The Patent Office

PCT/GB 93/02367

RECD 20 JAN 1994

WIPO PCT

08/486265

The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

I, the undersigned, being an officer duly authorised in accordance with Section 62(3) of the Patents and Designs Act 1907, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the Patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or the inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 7 January 1994.

PRIORITY DOCUMENT

9304677.9
08 Mar. 1993

For official use

THE PATENT OFFICE
- 8 MAR 1993
RECEIVED BY HAND

10MAR'93H00307730 PAT 1 77 UC 25.00

08 MAR 1993

9304677.9

Your reference 70/4201/01

Notes
Please type, or write in dark ink using CAPITAL letters. A prescribed fee is payable for a request for grant of a patent. For details, please contact the Patent Office (telephone 071–438 4700).

Rule 16 of the Patents Rules 1990 is the main rule governing the completion and filing of this form.

❷ Do not give trading styles, for example, 'Trading as XYZ company', nationality or former names, for example, 'formerly (known as) ABC Ltd' as these are not required.

Warning
After an application for a Patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977 and will inform the applicant if such prohibition or restriction is necessary. Applicants resident in the United Kingdom are also reminded that under Section 23, applications may not be filed abroad without written permission unless an application has been filed not less than 6 weeks previously in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction revoked.

The Patent Office

Request for grant of a Patent

Form 1/77

Patents Act 1977

❶ Title of invention

1 Please give the title of the invention: PROTEIN KINASES

❷ Applicant's details

❑ **First or only applicant**

2a If you are applying as a corporate body please give:

Corporate name: Ludwig Institute for Cancer Research

Country *(and State of incorporation, if appropriate)*: United Kingdom

2b If you are applying as an individual or one of a partnership please give in full:

Surname

Forenames

2c **In all cases**, please give the following details:

Address: St. Mary's Hospital Medical School
Norfolk Place
Paddington
London

UK postcode *(if applicable)*: W2 1PG

Country: United Kingdom

ADP number *(if known)*: 5356119001

KP

2d, 2e and 2f: If there are further applicants please provide details on a separate sheet of paper.

❑ **Second applicant *(if any)***

2d If you are applying as a corporate body please give:

Corporate name

Country *(and State of incorporation, if appropriate)*

2e If you are applying as an individual or one of a partnership please give in ful

Surname

Forenames

2f **In all cases**, please give the following details:

Address

UK postcode *(if applicable)*

Country

ADP number *(if known)*

❸ *An address for service in the United Kingdom must be supplied*

Please mark correct box

❸ Address for service details

3a Have you appointed an agent to deal with your application?

Yes [X] No [ ] ➡ ***go to 3b***

⬇ ***please give details below***

Agent's name GILL JENNINGS & EVERY

Agent's address* Broadgate House
7 Eldon Street
London

Postcode EC2M 7LH ✓

Agent's ADP number 745002

3b: If you have appointed an agent, all correspondence concerning your application will be sent to the agent's United Kingdom address.

3b If you have not appointed an agent please give a name and address in the United Kingdom to which all correspondence will be sent:

Name

Address

Postcode

Daytime telephone number *(if available)*

ADP number *(if known)*

eference number

4 Agent's or applicant's reference number *(if applicable)* 70/4201/01

❺ Claiming an earlier application date

5 Are you claiming that this application be treated as having been filed on the date of filing of an earlier application?

Please mark correct box

Yes ☐ No ☒ ➡ ***go to 6***

↓

please give details below

❑ number of earlier application or patent number

❑ filing date

(day month year)

❑ and the Section of the Patents Act 1977 under which you are claiming:

Please mark correct box

15(4) (Divisional) ☐ 8(3) ☐ 12(6) ☐ 37(4) ☐

❻ Declaration of priority

❻ If you are declaring priority from a PCT Application please enter 'PCT' as the country and enter the country code (for example, GB) as part of the application number.

Please give the date in all number format, for example, 31/05/90 for 31 May 1990.

6 If you are declaring priority from previous application(s), please give:

Country of filing	Priority application number *(if known)*	Filing date *(day, month, year)*

❼ *The answer must be 'No' if:*
- *any applicant is not an inventor*
- *there is an inventor who is not an applicant,* ***or***
- *any applicant is a corporate body.*

❼ Inventorship

7 Are you (the applicant or applicants) the sole inventor or the joint inventors?

Please mark correct box

Yes [ ] No [X] ➡ ***A Statement of Inventorship on Patents Form 7/77 will need to be filed (see Rule 15).***

❽ *Please supply duplicates of claim(s), abstract, description and drawing(s).*

❽ Checklist

8a Please fill in the number of sheets for each of the following types of document contained in this application.

Continuation sheets for this Patents Form 1/77			
Claim(s)		Description	27
Abstract		Drawing(s)	9

8b Which of the following documents also accompanies the application?

Please mark correct box(es)

Priority documents *(please state how many)*	
Translation(s) of Priority documents *(please state how many)*	
Patents Form 7/77 – Statement of Inventorship and Right to Grant *(please state how many)*	
Patents Form 9/77 – Preliminary Examination/Search	
Patents Form 10/77 – Request for Substantive Examination	

❾ *You or your appointed agent (see Rule 90 of the Patents Rules 1990) must sign this request.*

❾ Request

I/We request the grant of a patent on the basis of this application.

GILL JENNINGS & EVERY
Agents for the Applicant

Please sign here ➡ Signed [signature] Date 08. 03. 93 *(day month year)*

R E Perry

A completed fee sheet should preferably accompany the fee.

Please return the completed form, attachments and duplicates where requested, together with the prescribed fee to either:

❑ ***The Comptroller***
The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

or

❑ ***The Comptroller***
The Patent Office
25 Southampton Buildings
London
WC2A 1AY

PROTEIN KINASES

This invention is related to that described in Application No. 9224057.1, filed 17th November 1992. More particularly, it relates to

Human cDNA clones encoding four novel transmembrane protein serine/threonine kinases, denoted activin receptor-like kinase (ALK) -1, -2, -3 and -4, were obtained using a polymerase chain reaction (PCR)-based strategy. The PCR primers were designed based upon the sequence similarity between the activin receptor type II and Daf-1. The cDNA clones for ALK-1, -2 and -3 encode complete proteins of 503, 509 and 532 amino acids, respectively. The ALK-4 cDNA is incomplete and the predicted protein of 383 amino acids has a truncated extracellular domain. The ALKs share similar domain structures, comprising of predicted signal sequences at the N-terminals, followed by hydrophilic cysteine-rich binding domains, single hydrophobic transmembrane regions and C-terminal intracellular portions which consist almost entirely of putative serine/threonine kinase domains. The ALKs have approximately 40% sequence identity to activin receptors type II and IIB, TGF-β type II receptor and Daf-1 in the kinase domains. However, the sequence identities are higher (60-79%) between ALK-1, -2, -3 and -4, suggesting that they form a subfamily among the receptor serine/threonine kinases. The extracellular domains of ALKs show only little sequence identity to other receptor serine/threonine kinases, but the cysteine residues are conserved. The structural properties suggest that ALK-1 through 4 are receptors that may bind ligands that are members of the TGF-β superfamily. The expression of mRNA in human tissues varied for the different ALKs; ALK-2 and ALK-4 showed ubiquitous tissue expression patterns, whereas the distribution of ALK-1 and ALK-3 varied strongly between different tissues with more restricted expression patterns. These results suggest that each ALK may have different *in vivo* functions.

Introduction

Transforming growth factor-β (TGF-β) is a family of multifunctional peptides of which three isoforms, termed TGF-β1, TGF-β2 and TGF-β3, have been identified in mammals (for reviews see Roberts and Sporn, 1990; Moses *et al.*, 1990). The TGF-β family belongs to a larger superfamily of structurally related signalling proteins, including activins (Vale *et al.*, 1986; Ling *et al.*, 1986), inhibins (Mason *et al.*, 1985; Forage *et al.*, 1986), Müllerian inhibiting substance (MIS; Cate *et al.*, 1986), and bone morphogenetic proteins (BMPs; Wozney *et al.*, 1988; Wang *et al.*, 1988; Celeste *et al.*, 1990) or osteogenic proteins (OPs; Özkaynak *et al.*, 1990; 1992). All these molecules regulate cell proliferation, differentiation and various other cell functions (for reviews see Roberts and Sporn, 1990; Massagué, 1990).

Members of the TGF-β superfamily exert their effects on cells via binding to specific receptors. The receptors for TGF-β have been most thoroughly characterized (for review see Massagué, 1992; Lin and Lodish, 1993). By affinity labeling with radiolabeled TGF-β, a number of putative cell surface receptors have been identified, including three distinct size classes denoted receptor type I (53 kd), type II (75 kd), type III (or betaglycan, 300 kd proteoglycan with a 120 kd core protein) (Massagué and Like, 1985; Cheifetz *et al.*, 1986) and more recently endoglin (homodimer of two 95 kd subunits) (Cheifetz *et al.*, 1992). Current evidence implicates that receptors type I and type II are directly involved in receptor signal transduction (Segarini *et al.*, 1989; Boyd and Massagué, 1989; Laiho *et al.*, 1990); the receptors may interact and form a heteromeric complex (Laiho *et al.*, 1991). In fact, type I appears to require type II to bind TGF-β and type I appears to be necessary for type II to signal (Wrana *et al.*, 1992). Receptor type III and endoglin, which have recently been cloned (Wang *et al.*, 1991, López-Casillas *et al.*, 1991; Gougos and Letarte, 1990), may have more indirect roles in signalling, perhaps presenting ligands to signalling receptors.

Based upon the high structural similarity of the members of the TGF-β superfamily, the receptors for these ligands may have similar structural and functional characteristics. The TGF-β superfamily members do not compete with each other for receptor binding and thus appear to act through specific receptors to elicit different biological effects. In pituitary tumor cells, however, a distinct type IV TGF-β receptor, has been identified, which is also recognized by activin and inhibin (Cheifetz *et al.*, 1988). Binding studies with activin A and BMP-4 have shown two coexisting membrane components of 50-60 and 70-80 kd on responsive cells (Hino *et al.*, 1989; Paralkar *et al.*, 1991, Attisano *et al.*, 1992). In analogy with TGF-β receptors they may be signalling receptors and have been named receptors type I and type II. Direct evidence that receptors for TGF-β superfamily members indeed are structurally similar was provided by the recent expression cloning and sequence analysis of mouse activin receptor type II (mActR-II; Mathews and Vale, 1991) and human TGF-β receptor type II (hTβR-II; Lin *et al.*, 1992). These genes show sequence homology with each other and with a previously identified *C. elegans* putative receptor gene *daf-1* that controls Dauer Larva development (Georgi *et al.*, 1990). Additionally, another structurally related activin receptor, type IIB (ActR-IIB), was identified using a low stringency hybridization protocol (Mathews *et al.*, 1992), and using a PCR-based approach (Attisano *et al.*, 1992; Legerski *et al.*, 1992). The encoded proteins are all putative serine/threonine kinase receptors, providing a possible mechanism for transmembrane signalling.

The kinase domains of the serine/threonine kinase receptors show significant sequence similarity which define them as a separate kinase family. Autophosphorylation on serine and threonine residues of the kinase domain of hTβR-II has been demonstrated in a bacterial fusion protein (Lin *et al.*, 1992). The extracellular domains of these receptors show little sequence similarity, but all are rich in cysteine residues (Lin *et al.*, 1992).

To identify additional members of the receptor serine/threonine kinase family, which may bind members of TGF-β superfamily, we have used a PCR-based approach. The strategy made use of the sequence identity between short stretches of conserved amino acids, in particular in the kinase domains of mActR-II and Daf-1, and resulted in the isolation of four novel members of a subfamily of receptor serine/threonine kinases, denoted activin receptor-like kinases (ALKs).

Results

Identification of novel genes that encode receptor serine/threonine kinases

To identify novel receptor serine/threonine kinases we used a PCR based approach similar to that used for the identification and cloning of novel tyrosine kinases (Wilks, 1989; Kamb *et al.*, 1989; Reid *et al.*, 1990; Partanen *et al.*, 1990; 1992; Raz *et al.*, 1991; Masiakowski and Carrol, 1992). HEL cells were chosen as sources of mRNA in reverse transcription (RT)-PCR, as they have been shown to respond to both activin and TGF-β (Hino *et al.*, 1989; Sing *et al.*, 1988).

Degenerate PCR primers were designed based upon short stretches of amino acid sequence identity in the mActR-II (Mathews and Vale, 1991) and *daf-1* gene product (Georgi *et al.*, 1990). Three PCR primers allowed us to obtain PCR recombinants encoding novel amino acid sequences with conservation of specific amino acid residues found in ActR-II and Daf-1; primer B3-S, derived from the kinase subdomain II (nomenclature according to Hanks *et al.*, 1988), encodes the lysine residue involved in ATP binding, and primers B7-S and E8-AS, derived from the kinase subdomains VII and VIII, encode serine/theronine kinase specific residues.

Using primers B3-S and E8-AS, the PCR recombinants derived from previously identified ActR-II and TβR-II as well as three novel PCR recombinants were identified, termed 11.1, 11.2 and 11.3, and using primers B7-S and E8-AS, four other novel PCR recombinants were identified, including PCR recombinant 5.2. The four receptor genes that are presented in this report correspond to four different PCR recombinants that were closely related to the activin and TGF-β type II receptors.

cDNA cloning and sequence analysis of ALK-1, ALK-2, ALK-3, and ALK-4

To clone cDNA for ALK-1 we screened an oligo(dT) primed human placenta cDNA library with a radiolabeled insert derived from the PCR recombinant 11.3. Three different types of clones were identified, based upon their *Eco*RI restriction enzyme pattern. Many positive clones of one class with an approximate insert size of 2 kb, were isolated in the library screening, which are likely to correspond to the 2.2 kb transcript size seen in Northern blots probed for ALK-1 (see Figure 4A). The cDNA clone HP 57 was chosen as representative of this class and subjected to complete sequencing. Of the other two types only one clone was isolated with insert sizes of 1.7 kb and 3.5 kb, respectively. The latter cDNA clone may correspond to the longer ALK-1 transcript seen in Northern blots (data not shown).

Sequence analysis of ALK-1 (cDNA clone HP 57) revealed a sequence of 1984 nucleotides including a poly A tail (Figure 1A). The longest open reading frame encodes a protein of 503 amino acids, with high sequence identity to receptor serine/threonine kinases (see below). The first methionine codon, the putative translation start site, is at nucleotides 283-285 and is preceeded by an in-frame stop codon. This first ATG is in a more favorable context for translation initiation (Kozak, 1987) than the second and third in-frame ATG at nucleotides 316-318 and 325-327. The putative initiation codon

is preceeded by a 5' untranslated sequence of 282 nucleotides that is GC-rich (80 % GC), which is not uncommon for growth factor receptors (Kozak, 1991). The 3' untranslated sequence comprises 193 nucleotides and ends with a poly A tail. No *bona fide* poly A addition signal is found, but there is a sequence AATACA 17-22 nucleotides upstream of the poly A tail which could possibly serve as a poly A addition signal.

ALK-2 cDNA was cloned by screening an amplified oligo(dT) primed human placenta cDNA library with a radiolabeled insert derived from the recombinant 11.2. Two clones termed HP 53 and HP64 with insert sizes of 2.7 kb and 2.4 kb, respectively, were identified, and their sequences were determined. We found no sequence difference in the overlapping clones and conclude that both are derived from transcripts of the same gene.

Sequence analysis of cDNA clone HP 53 for ALK-2 (Figure 1B) revealed a sequence of 2719 nucleotides with a poly A tail. The longest open reading frame encodes a protein of 509 amino acids. The first ATG at nucleotides 104-106 agrees favorable with Kozak's consensus sequence with an A at position -3. This ATG is preceded in-frame by a stop codon. There are four ATG codons in close proximity further downstream, which agree with the Kozak's consensus sequence (Kozak, 1987), but according to Kozak's scanning model the first ATG is predicted to be the translation start site (Kozak, 1987). The 5' untranslated sequence is 103 nucleotides. The 3' untranslated sequence of 1089 nucleotides contains a polyadenylation signal located 9-14 nucleotides upstream from the poly A tail. The cDNA clone HP 64 lacks 498 nucleotides from 5' end compared to HP 53, but the sequence extended at the 3' end with 190 nucleotides and poly A tail is absent (data not shown). This suggests that different polyadenylation sites occur for ALK-2. In Northern blots, however, only one transcript was detected (see below).

The cDNA for ALK-3 was cloned by initially screening an oligo(dT) primed human foreskin fibroblast cDNA library with an oligonucleotide derived

from the PCR recombinant 5.2. One positive cDNA clone with an insert size of 3 kb, ON11, was identified. Upon partial sequencing, it appeared that this clone was incomplete; it encodes only part of the kinase domain and lacks the extracellular domain. Sequences corresponding to the truncated kinase domain were subsequently used to probe a random primed fibroblast cDNA library from which one cDNA clone with an insert size of 3 kb, termed ONF5 (Figure 1C), was isolated. Sequence analysis of ONF5 revealed a sequence of 2932 nucleotides without poly A tail, suggesting that this clone was derived by internal priming. The longest open reading frame codes for a protein of 532 amino acids. The first ATG codon is compatible with Kozak's consensus sequence (Kozak, 1987), is at 310-312 nucleotides and is preceeded by an in-frame stop codon. The 5' and 3' untranslated sequences are 309 and 1027 nucleotides long, respectively.

ALK-4 cDNA was identified by screening a human oligo(dT) primed human erythroleukemia cDNA library with the PCR recombinant 11.1 as a probe. One cDNA clone, termed 11H8, was identified with an insert size of 2 kb (Figure 1D). An open reading frame was found encoding a protein sequence of 383 amino acids with high similarity to receptor serine/threonine kinases; however, the extracellular domain was truncated, indicating that the isolated cDNA was incomplete. The 3' untranslated sequence is 818 nucleotides and does not contain a poly A tail, suggesting that the cDNA was internally primed.

Comparison of the structures of ALKs

ALK-1, ALK-2 and ALK-3 have a similar domain structure (Figure 2). Hydrophilic, relatively short (97 to 129 amino acids) putative extracellular domains are flanked by hydrophobic leader sequences at the N-terminals (von Heijne, 1986), and by putative single transmembrane domains (Kyte and Doolittle, 1982), which are followed by cytoplasmic domains at the C-

terminals. The putative ligand binding domains of ALKs share little sequence identity (15-20%) but each contains 10 conserved cysteine residues (Figure 2). Thus, it is likely that their three dimensional structures are similar. Each of the ALKs has one potential N-linked glycosylation site; its position is conserved in ALK-1 and ALK-2, but not in ALK-3. The intracellular portions of ALK-1, -2, -3 and -4 consist almost entirely of kinase domains that have 60-79% sequence identity (Figure 2 and 3A). Based upon the sequence similarity between activin and TGF-β type II receptors, the C-termini of the kinase domains of ALK-1, ALK-2, ALK-3 and ALK-4 are set at Ser-495, Ser-501, Ser-527 and Gln-378, respectively. A consensus sequence for the binding of ATP (Gly-X-X-Gly-X-X-Gly followed by Lys further downstream in subdomain II) is found in four ALKs. Analysis of the amino acid sequences in subdomains VI and VIII, that are most useful to distinguish the specificity of amino acid phosphorylation, predicts that ALKs are serine/threonine kinases (see discussion). Two kinase inserts are present between kinase subdomains VIA and VIB, and between subdomains X and XI, respectively (Figure 2). The regions of highest divergence in the intracellular parts of ALKs are in the juxtamembrane segments (Figure 2).

Phylogenetic relationship of receptor serine/threonine kinases

A search of the individual ALKs in the database identifies ActR-II, ActR-IIB, TβR-II and Daf-1 as the closest homologues of ALKs, with which they share approximately 40% amino acid identity (Figure 3A).

The sequence of an additional member of the receptor serine/threonine kinase family, ALK-5, is presented elsewhere (Franzén *et al.*, submitted for publication). ALK-5 and ALK-4 have 90% sequence identity in their kinase domains (Figure 3A). In comparison, ActR-II and ActR-IIB, which both bind activin, have a sequence identity in their kinase domains of 78%. Pairwise comparison of members of the serine/threonine kinase family of receptors

using the Jutun-Hein sequence alignment program (Hein, 1990), identifies ALK-1 through -5 as a separate subclass among the receptor serine/threonine kinases (Figure 3B).

mRNA expression of ALK-1 through -4

The distribution of ALK-1, -2, -3 and -4 in various human tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas) was determined by Northern blot analysis of mRNA (Figure 4). In order to minimize cross-hybridization, probes were used that did not encode part of the kinase domains, but corresponded to the highly diverged sequences of either 5' untranslated and ligand binding regions (probes for ALK-1, -2, and -3) or 3' untranslated sequences (probe for ALK-4).

Using the probe for ALK-1, two trancripts of 2.2 and 4.9 kb were detected (Figure 4A). The ALK-1 expression level varied strongly between different tissues; high in placenta and lung, moderate in heart, muscle and kidney, and low (to not detectable) in brain, liver and pancreas. The relative ratios between the two transcripts were similar in most tissues; in kidney, however, there was relatively more of the 4.9 kb transcript. By reprobing the blot with mRNA from different tissues using a probe for ALK-2, one transcript of 4.0 kb was detected (Figure 4B), with a ubiquitous expression pattern. Expression was detected in every tissue investigated and was highest in placenta and skeletal muscle. Subsequently, the multiple tissue blot was reprobed with a probe for ALK-3. One major transcript of 4.4 kb and a minor transcript of 7.9 kb were detected (Figure 4C). Expression was high in skeletal muscle, in which also an additional minor transcript of 10 kb was observed. Moderate levels of ALK-3 mRNA were detected in heart, placenta, kidney and pancreas, and low (to not detectable) expression was found in brain, lung and liver. The relative ratios between the different trancripts were similar in the tested tissues with the 4.4 kb transcript being the predominant one, with the

exception for brain where both transcripts were expressed at a similar level. Probing the blot with ALK-4 indicated the presence of a transcript with the estimated size of 5.2 kb, and revealed an ubiquitous expression pattern (Figure 4D).

Discussion

Cellular proliferation and differentiation are controlled, in part, by a complex network of interacting growth factors. A large number of growth factors mediate their effects by binding to cell surface receptors that possess intrinsic ligand-activated tyrosine kinase activities (for review see, Ullrich and Schlessinger, 1990). Another family of transmembrane receptors for growth regulatory factors with serine/threonine kinase domains has recently been found and is rapidly growing. The present study was initiated to clone additional members of this receptor family that may bind ligands of the TGF-β superfamily. Here we describe the cloning and sequence analysis of four novel human receptors, ALK-1, -2, -3, and -4, which together with ALK-5 (Franzén et al., submitted), ActR-II, ActR-IIB and TβR-II, constitute the mammalian receptor serine/threonine kinase family.

The ALKs share a similar domain structure, i.e. relatively short extracellular domains, transmembrane domains and cytoplasmic protein kinase domains. The extracellular domains share little sequence identity, but contain conserved cysteine residues. In the intracellular portion the kinase domain of ALKs share a high (60-79%) sequence identity and juxtamembrane segments are most dissimilar. The significantly higher sequence identities between the kinase domains of ALKs compared with the identities between ALKs and other kinases defines ALKs as a separate subfamily among the serine/threonine kinases. A characteristic for ALKs, shared with ActR-II, TβR-II and Daf-1, is that the protein kinase domain is interrupted by two short kinase

inserts between subdomains VIA and VIB, and subdomains X and XI. ActR-II, TβR-II and Daf-1 have C-terminal tails of 23-28 amino acids that are rich in serine and threonine residues and could function as autophosphorylation sites. ALKs, however, have relatively short C-terminal tails of only 5 to 8 amino acid residues, which lack serine and threonine residues exept for one threonine residue in ALK-2. Comparison of ALKs with other mammalian receptor serine/threonine kinases shows that the regions that diverge most in the intracellular domains are in the juxtamembrane segments, kinase inserts and C-terminal tails. If these regions, in analogy with the situation that prevail for tyrosine kinases, are important for interaction with down-stream components in the signal transduction pathway, then ALKs as a group may share such components, but the specificity is likely to be different for other receptor serine/threonine kinases.

The residues of ALKs in the kinase subdomain VIB and VIII predict that they are serine/threonine kinases. The sequence motifs HRDLKSKN (subdomain VIB) and GTKRYMAPE (subdomain VIII) are highly conserved in ALKs, ActR-II, ActR-IIB, TβR-II and Daf-1 and other serine/threonine kinases, but not in tyrosine kinases. In tyrosine kinases tyrosine residue(s) surrounded by acidic amino acids are found between subdomains VII and VIII. In ALK-1, but not in other ALKs, a tyrosine residue is present at position 361 and is flanked by two aspartic acid residues. Of note, purified mActR-IIB was recently shown to be a dual kinase which phosphorylates itself as well as exogenous substrates on serine, threonine and tyrosine residues (Nakamura *et al.*, 1992). Our current work is aimed at verifying that ALKs have kinase activity, and determine whether serine, threonine or tyrosine residues are phosphorylated.

The high sequence identity between ALKs and ActR-II and TβR-II suggests that ALKs may also function as signalling type II receptors for members of the TGF-β superfamily. However, the molecular masses of the

mature ALK-1, -2 and -3 proteins are 53.9 kd, 55.7 kd and 57.5 kd, respectively; thus the possibility that members of the ALK family are type I receptors for TGF-ß superfamily member, has not been excluded.

Whereas the expressions of ALK-2, ALK-4 and ALK-5 are ubiquitous, the expressions of ALK-1 and ALK-3 show more restricted patterns (Figure 4 and Franzén *et al.*, manuscript submitted). The expression profile of each ALKs is different. Thus, it is likely that the different ALKs serve different functions. The accumulated data support the notion that growth regulatory factors act preferentially in autocrine and/or paracrine manners. If this is correct for the ligands for ALKs, ALK-2 and ALK-4 may bind to ligands with effects on many tissue types, exemplified by TGF-β. ALK-1, predominantly expressed in placenta and lung, and ALK-3, highly expressed in muscle, may bind to ligands with more restricted action range, exemplified by MIS. Further studies are in progress to identify the ligands for ALKs. This involves a detailed characterization of the expression patterns of ALKs to allow a comparison with those of possible ligands, as well as direct binding experiments.

All detected transcript sizes were different, and thus no cross-reaction between mRNAs for the different ALKs was observed when the specific probes were used. This suggests that the multiple transcripts of ALK-1 and ALK-3 are coded from the same gene. The mechanism for generation of the different transcripts is unknown at present; they may be formed by alternative mRNA splicing, different polyadenylation, use of different promoters, or by a combination of these events. The functional significance of the presence of different transcripts also remains to be elucidated. Difference in mRNA splicing in the regions coding for the extracellular domains may lead to the synthesis of receptors with different affinities for ligands, as was shown for mActR-IIB (Attisano *et al.*, 1992), or the production of soluble binding proteins. Differences in the intracellular domains may lead to truncated cell surface

receptors that may possibly act in a dominant negative fashion, or act as receptors with different enzymatic activity or substrate specificity.

Many TGF-β superfamily members and their receptors appear to have important functions during early development in mediating the induction and pattern formation of mesodermal tissues. The injection of mRNA for ActR-II (Kondo *et al.*, 1991) and ActR-IIB (Mathews *et al.*, 1992) leads to embryos that display developmental defects characterized by inappropriate formation of dorsal mesodermal tissue. ALKs may also have important roles during the development. Thus, it will be important to determine their expression in different tissues during different developmental stages, and to measure their effects on mesoderm formation after injection of their mRNA into *Xenopus* embryos.

Our experimental strategy appears to have selectively amplified sequences encoding parts of putative receptor serine/threonine kinase genes. Thus far no genes which encode cytoplasmic serine/threonine kinases were picked up. Since about 20 members of TGF-β superfamily have been identified, it is anticipated that there are more members of the receptor serine/threonine kinase family. It is also possible that some of the serine/threonine kinases may bind other proteins unrelated to the members of the TGF-β superfamily, e.g. members of Wnt protein family. The novel genes we identified by PCR were all derived from mRNA from HEL cells. It is likely that use of other mRNA sources, in combination with the same or similar PCR primers, will lead to the identification of additional members in this receptor family.

Materials and methods

Nucleic acids

RNA was prepared from HEL cells using the guanidium isothiocyanate method (Chirgwin *et al.*, 1979). Poly(A+) RNA was selected by polyAT tract system (Promega) as described by the manufacturers or purified through oligo(dT)-cellulose column as described (Aviv and Leder, 1972). Oligonucleotides were synthesized with a Gene assembler plus (Pharmacia-LKB) using phosphoramidite technique, deprotected in 30% ammonia for 2 hr at 80°C, and then desalted on NAP-10 columns (Pharmacia-LKB).

RT-PCR

2 µg of HEL mRNA was reverse transcribed into cDNA in the presence of 50 mM Tris-HCl, pH 8.3, 8 mM magnesium chloride, 30 mM potassium cloride, 10 mM dithiothreitol, 2 mM nucleotide triphosphates, excess oligo(dT) primers and 34 units of AMV reverse transcriptase at 42°C for 2 hr in 40 µl of reaction volume. For amplification of the mActR-II/Daf-1 like sequences three different primers were used. The primers were designed so that each primer set had a similar GC content, and so that self complementarity and complementarity between the 3' ends of the primers were avoided. The degeneracy of the primers was kept as low as possible, by avoiding serine, leucine and arginine residues, which have 6 possible codons, and applying human codon preference. Degeneracy was particularly avoided at the 3' end, since unlike the 5' end where mismatches are tolerated, mismatches at the 3' end may dramatically reduce the efficiency of PCR. In order to facilitate directional subcloning, sequences were included at the 5' end that lead to restriction sites in the PCR products; GC clamps were then included at the most 5' end to permit efficient restriction enzyme digestion. The B3-S primer is a 25-mer sense oligonucleotide containing a *Bam*HI site at a 5' end with a 162-fold degeneracy (5'- GCGGATCCGT(CGT)GC(ACT)GT(CGT)AA(AG)AT (TCA)TT-3'), derived from the conserved sequence motif "VAVKIF" in the kinase subdomain II. The B7-S primer is a 24-mer sense oligonucleotide

containing a *Bam*HI site at the 5' end with a 288-fold degeneracy (5'-GCGGATCCG(CAG)GA(TC)AT(TCA)AA(AG)(AT)(GC)(CT)AA-3'), derived from the "RDIKSKN" sequence motif in the kinase subdomain VII. The E8-AS primer is a 20-mer antisense oligonucleotide containing an *Eco*RI site at the 5' end with a degeneracy of 18-fold (5'-CGGAATTC(TGA)GG(TGA)GCCAT (AG)TA-3'), derived from the "YMAPE" sequence motif from the kinase domain VIII. Amplification by PCR was carried out with a 7.5% aliquot (3 µl) of the reverse transcribed mRNA, in the presence of 10 mM Tris-HCl, pH 8.3, 50 mM potassium chloride, 1.5 mM magnesium chloride, 0.01% gelatin, 0.2 mM nucleotide triphosphates, 1 µM of both sense and antisense primers and 2.5 units of Taq polymerase (Perkin Elmer Cetus) in 100 µl reaction volume. Amplifications were performed on a thermal cycler (Perkin Elmer Cetus) using the following program: first 5 cycles with denaturation for 1 min at 94°C, annealing for 1 min at 50°C, a 2 min ramp to 55°C and elongation for 1 min at 72°C, followed by 20 cycles of 1 min at 94°C, 30 sec at 55°C and 1 min at 72°C. Thereafter, a 3% aliquot (3 µl) of the material of the first reaction was taken and processed for a second round of PCR amplification; 25 cycles of 1 min at 94°C, 30 sec at 55°C, and 1 min at 72°C. Samples were digested with *Bam*HI and *Eco*RI and subsequently fractionated by low melting point agarose gel electrophoresis. Bands corresponding to the approximate expected sizes (460 bp for primer pair B3-S and E8-AS and 140 bp for primer pair B7-S and E8-AS) were excized from the gel and DNA was purified. Subsequently, these fragments were ligated into pUC19, that had been previously linearized with *Bam*HI and *Eco*RI and transformed into *E. coli* strain DH5α using standard protocols (Sambrook *et al.*, 1989).

cDNA libraries

Poly(A+) RNA was isolated from human term placenta. Subsequently, the poly(A+)RNA was converted into cDNA using MLV reverse transcriptase

(Superscript, BRL) and used to generate an oligo(dT) primed human placenta λZAP II cDNA library of 5 x 10^5 independent clones, which was amplified. Poly(A+) RNA isolated from AG 1518 human foreskin fibroblasts was used to prepare a primary random primed λZAP II cDNA library of 1.5 x 10^6 independent clones using Ribo Clone cDNA synthesis system (Promega) and Gigapack Gold II packaging extract (Statagene). A primary oligo(dT) primed human foreskin fibroblast λgt10 cDNA library (Claesson-Welsh *et al.*, 1989) was generously provided by Lena Claesson-Welsh. An amplified oligo(dT) primed HEL cell λgt11 cDNA library of 1.5 x 10^6 independent clones (Poncz *et al.*, 1987) was generously provided by Mortimer Poncz and Kari Alitalo.

Isolation of cDNA clones

Employing RT-PCR on HEL mRNA with the primer pair B3-S and E8-AS, three PCR recombinants were obtained, termed 11.1, 11.2 and 11.3, that corresponded to novel genes. Using the primer pair B7-S and E8-AS, four additional novel PCR recombinants were obtained of which one termed PCR recombinant 5.2 is presented in this report. For ALK-1, the placenta cDNA library was screened with radiolabeled insert of PCR recombinant 11.3. Based upon their restriction enzyme digestion pattern, three different types of clones with approximate insert sizes of 1.7 kb, 2 kb and 3.5 kb were identified. Clone HP 57 with an insert size of 2 kb was chosen for complete sequencing. For ALK-2 the placenta library was screened with the PCR product 11.2. Two clones, termed HP53 and HP64 were identified with insert sizes of 2.7 kb and 2.4 kb, respectively, and their DNA sequences were determined. For ALK-3, the oligo(dT) primed fibroblast cDNA library was screened with an oligonucleotide (5'-ATTCAAGGGCACATCAACTTCATTTGTGTCACTGTTG-3') derived from the PCR recombinant 5.2. One partial clone, termed ON11, with an insert size of 3 kb, was isolated and partially sequenced. The most 5' sequences, a 540 nucleotides *Xba*I restriction fragment encoding a truncated

kinase domain, was used to screen the random primed fibroblast cDNA library. One clone, ONF5, with an insert size of 3 kb was obtained, and its sequence was determined. For ALK-4, the oligo(dT) primed HEL cell cDNA library was screened with the radiolabeled insert of the PCR recombinant 11.1 as a probe. One positive clone with an insert size of 2 kb was obtained and its sequence was determined. Hybridization and purification of positive bacteriophages were performed as described previously (Kanzaki *et al.*, 1990). Labeling of the inserts of PCR recombinants was performed using random priming method (Feinberg and Vogelstein, 1983) and oligonucleotide derived from the sequence of the PCR recombinant 5.2 was labeled by phosphorylation with T4 polynucleotide kinase following standard protocols (Sambrook *et al.*, 1989).

DNA sequencing and sequence analysis

Double stranded DNA was sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977) using T7 DNA polymerase (Pharmacia-LKB). Sequencing of plasmid inserts was performed using M13 universal, M13 reverse as well as specific primers and unidirectional deletion clones generated by Exonuclease III digestion (Henikoff,1984). Compressions were resolved using 7-deaza-GTP (U.S. Biochemical Corp.). DNA sequences were analysed using DNA STAR computer program (DNA STAR).

Northern blot analysis

RNA blots were hybridized with ^{32}P-labeled probes at 42°C overnight in 50% formamide, 5 x standard saline citrate (SSC; 1 x SSC is 50 mM sodium citrate, pH 7.0, 150 mM sodium chloride), 0.1% sodium dodecyl sulfate (SDS), 50 mM sodium phosphate and 0.1 mg/ml salmon sperm DNA. The probes were labeled by random priming (Feinberg and Vogelstein, 1983) and unincorporated label was removed by Sephadex G-25 chromatography.

Filters were washed at 65°C, two times for 30 min in 2.5 x SSC, 0.1% SDS and two times for 30 min in 0.3 x SSC , 0.1% SDS. Stripping of the blots was performed by incubation at 90°C - 100°C in water for 20 min.

Acknowledgments

We thank Christer Wernstedt for synthesis of oligonucleotides, Susanne Grimsby for technical assistance, and Lena Claesson-Welsh, Mortimer Poncz and Kari Alitalo for providing cDNA libraries. Peter ten Dijke is supported by an EMBO postdoctoral fellowship.

References

Attisano, L., Wrana, J.L., Cheifetz, S. and Massagué, J. (1992) *Cell*, **68**, 97-108.

Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408-1412.

Boyd, F.T. and Massagué, J. (1989) *J. Biol. Chem.*, **264**, 2272-2278.

Cate, R.L., Mattaliano, R.J., Hession, C.. Tizard, R., Farber, N.M., Cheung, A., Ninfa, E.G., Frey, A. Z., Gash, D.J., Chow, E.P., Fisher, R.A., Bertonis, J.M., Torres, G., Wallner, B.P., Ramachandran, K.L., Ragin, R.C., Manganaro, T.F., MacLaughlin, D.T. and Donahoe, P.K. (1986) *Cell*, **45**, 685-698.

Celeste, A.J., Iannazzi, J.A., Taylor, R.C., Hewick, R.M., Rosen, V., Wang, E.A. and Wozney, J.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 9843-9847.

Cheifetz, S., Bellon, T, Calés, C., Vera, S., Bernabeu, C., Massagué, J. and Letarte, M. (1992) *J. Biol. Chem.*, **267**, 19027-19030.

Cheifetz, S., Like, B. and Massagué, J. (1986) *J. Biol. Chem.*, **261**, 9972-9978.

Cheifetz, S , Ling, N., Guillemin, R., and Massagué, J.(1988) *J. Biol. Chem.* **263**, 17225-17228.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry*, **18**, 5294-5299.

Claesson-Welsh, L., Eriksson, A., Westermark, B. and Heldin, C.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4917-4921.

Donaldson, C.J., Mathews, L.S. and Vale, W.W. (1992) *Biochem. Biophys. Res. Commun.* **184**, 310-316.

Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.*, **132**, 6-13.

Forage R.G., Ring, J.M., Brown, R.W., McInerney, B.V., Cobon, G.S., Gregson, R.P., Robertson, D.M., Morgan, F.J., Hearn, M.T. and Findlay, J.K. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3091-3095. (1986)

Georgi, L.L., Albert, P.S. and Riddle, D.L. (1990) *Cell*, **61**, 635-645.

Gougos A. and Letarte, M. (1990) *J. Biol. Chem.*, **265**, 8361-8364.

Hanks, S. E., Quinn, A.M. and Hunter, T. (1988) *Science*, **241**, 42-52.

Hein, J.J. (1990) *Meth. Enzymol.* **183**, 626-645.

Henikoff, S. (1984) *Gene*, **28**, 351-359.

Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* , **5**, 151-153.

Hino, M., Tojo, A., Miyazono, K., Miura, Y., Chiba, S., Eto, Y., Shibai, H. and Takaku, F. (1989) *J. Biol. Chem.* **264**, 10309-10314.

Kamb, A., Weir, M., Rudy, B., Varmus, H. and Canypn, C. (1989) *Proc. Natl. Acad. Sci USA*, **86**, 4372-4376.

Kanzaki, T., Olofsson, A., Morén, A., Wernstedt, C., Hellman, U., Miyazono, K., Claesson-Welsh, L. and Heldin, C.-H. (1990) *Cell*, **61**, 1051-1061.

Kondo, M., Tashiro, K., Fujii, G., Asano, M., Miyoshi, R., Yamada, R., Muramatsu, M. and Shiokawa, K. (1991) *Biochem. Biophys. Res. Commun.*, **181**, 684-690.

Kozak, M. (1987) *Nucl. Acids Res.*, **15**, 8125-8148.

Kozak, M. (1991) *J. Cell. Biol.*, **115**, 887-903.

Kyte, J. and Doolittle R.F. (1982) *J. Mol. Biol.*, **157**, 105-132.

Laiho, M., Weis, F.M.B. and Massagué, J. (1990) *J. Biol. Chem.*, **265**, 18518-18524.

Laiho, M., Weis, F.M.B., Boyd, F.T., Ignotz, R.A. and Massagué, J. (1991) *J. Biol. Chem.*, **266**, 9108-9112.

Legerski, R., Zhou, X., Dresback, J., Eberspaecher, H., McKinney, S., Segarini, P. and de Crombrugghe, B. (1992) *Biochem. Biophys. Res. Commun.*, **183**, 672-679.

Lin, H.Y., Wang, X.-F., Ng-Eaton, E., Weinberg, R.A. and Lodish, H.F. (1992) *Cell*, **68**, 775-785.

Lin, H.Y. and Lodish, H.F (1993) *Trends Cell Biol.*, **3**, 14-19.

Ling, N., Ying, S.Y., Ueno, N., Shimasaki, S., Esch, F., Hotta, M. and Guillemin, R. (1986) *Nature,* **321**, 779-782.

López-Casillas, F., Cheifetz, S., Doody, J., Andres, J.L., Lane, W.S. and Massagué, J. (1991) *Cell,* **67**, 785-795.

Masiakowski, P. and Carrol, R.D. (1992) *J. Biol. Chem.,* **267**, 26181-26190.

Mason, A.J., Hayflick, J.S., Ling, N., Esch, F., Ueno, N, Ying, S.-Y., Guillemin, R., Niall, H., and Seeburg, P.H. (1985) *Nature,* **318**, 659-663.

Massagué, J. (1990) *Annu. Rev. Cell Biol.,* **6**, 597-641.

Massagué, J. (1992) *Cell,* **69**, 1067-1070.

Massagué, J. and Like, B. (1985) *J. Biol. Chem.*, **260**, 2636-2645.

Mathews, L.S. and Vale, W.W. (1991) *Cell,* **65**, 973-982.

Mathews, L.S., Vale, W.W. and Kintner, C.R. (1992) *Science,* 255, 1702-1705.

Matzuk, M.M. and Bradley, A. (1992) *Biochim. Biophys. Acta*, **1130**, 105-108.

Moses, H.L., Yang, E.Y. and Pietenpol, J.A. (1990) *Cell,* **63**, 245-247.

Nakamura, T., Sugino, K., Kurosawa, N., Takio, K., Eto, Y., Iwashita, S., Muramatsu, M. Titani, K. and Sugino, H. (1992) *J. Biol. Chem.,* **267**, 18924-18928.

Özkaynak, E., Rueger, D.C., Drier, E.A., Corbett, C., Ridge, R.J., Sampath, T.K. and Opperman, H. (1990) *EMBO J.*, **9**, 2085-2093.

Özkaynak, E., Schnegelsberg, P.N.J., Jin, D.F., Clifford, G.M., Warren, F.D., Drier, E. A. and Opperman, H. (1992) *J. Biol. Chem.*, **267**, 25220-25227.

Paralkar, V. M., Hammonds, R.G. and Reddi. A.H. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3397-3401.

Partanen, J., Armstrong, E., Mäkelä, T.P., Korhonen, J., Sandberg, M., Renkonen, R., Knuutila, S., Huebner, K. and Alitalo, K. (1992) *Mol. Cell. Biol.,* **12**, 1698-1707.

Partanen, J., Mäkelä, T.P., Alitalo, R., Lehväslaiho, H. and Alitalo, K. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8913-8917.

Poncz, M. Surrey, S., LaRocco, P., Weiss, M., Rappaport, E.F., Conway, T.M. and Schwartz, E. (1987) *Blood*, **69**, 219-223.

Raz, V., Kelman, Z., Avivi, A., Neufeld, G., Givol, D. and Yarden, Y. (1991) *Oncogene*, **6**, 753-760.

Reid, H.H., Wilks A.F. and Bernards, O. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1596-1600.

Roberts, A. B. and Sporn, M. B. (1990) In Sporn, M.B. and Roberts, A.B. (eds.) *Peptide Growth Factors and Their Receptors, Part I*, Springer-Verlag, Berlin, pp. 419-472.

Sambrook, J., Fritsch, E.F.. and Maniatis, T.. (1989) *Molecular cloning: A laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

Sanger, R., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Segarini, P.R., Rosen, D.M. and Seyedin, S.M. (1989) *Mol. Endocrinol.*, **3**, 261-272.

Sing, G.K., Keller, J.R., Ellingsworth, L.R. and Ruscetti, F.W. (1988) *Blood* **72**, 1504-1511.

Ullrich, A., and Schlessinger, J. (1990) *Cell*, **61**, 203-212.

Vale, W.W., Rivier, J., Vaughan, J., McClintock, R., Corrigan, A., Woo, W., Karr, D. and Spiess, J. (1986) *Nature*, **321**, 776-779.

von Heijne, G. (1986) *Nucl. Acids Res.*, **14**, 4683-4690.

Wang, X.-F., Lin, H.Y., Ng-Eaton, E., Downward, J., Lodish, H.F. and Weinberg, R.A. (1991) *Cell*, **67**, 797-805.

Wang E.A., Rosen, V., Cordes, P., Hewick, R.M., Kriz, M.J., Luxenberg, D.P., Sibley, B.S., and Wozney, J.M. (1988) *Pcoc. Natl. Acad. Sci. USA*, **85**, 9484-9488.

Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1603-1607.

Wozney, J.M., Rosen, V., Celeste. A.J., Mitsock, L.M., Whitters, M.J., Kriz, R.W., Hewick, R.M. and Wang, E.A. (1988) *Science*, **242**, 1528-1534.

Wrana, J.L., Attisano, L., Cárcamo, J., Zentella, A., Doody, J. and Massagué, J. (1992) *Cell*, **71**, 1003-1014.

Figure legends

Fig. 1. Nucleotide and deduced amino acid sequences of cDNA's for ALK-1 (A), ALK-2 (B), ALK-3 (C) and ALK-4 (D). For each of the ALKs the following features are indicated; the N-terminal hydrophobic signal sequences as determined by the von Heijne algorithm (von Heijne, 1986) and the putative transmembrane sequences are overlined (thin lines); the sites for potential N-linked glycosylation are overlined (thick lines), cysteine residues found in the extracellular domain are boxed and the borders of the kinase domains are marked by arrows. The regions of the primers that lead to the identification of the sequences are indicated by half arrows. The stop codons which end the open reading frame are marked by asterisks. The putative polyadenylation signal AATAAA, if present, is underlined. The nucleotide sequences are numbered from the 5' ends of the cDNA clones and protein sequences are numbered starting from the N-termini. Both are indicated at the right of each lane.

Fig. 2. Comparison of human amino acid sequences of ActR-II (Donaldson *et al.*, 1992; Matzuk and Bradley, 1992), TβR-II (Lin *et al.*, 1992) and ALK-1 through -4. Identical amino acids are boxed and conserved cysteine residues in the extracellular domains are indicated with dots. Putative transmembrane regions are double overlined. The borders of the kinase domains are indicated by arrows. Kinase subdomains are indicated with roman numerals, according to the nomenclature of Hanks *et al.* (1988). Alignment was performed using the Clustal V alignment program (Higgins and Sharp, 1989), with some manual adjustment.

Fig. 3. Comparison of kinase domains of serine/threonine kinases. (A) Percent amino acid identity of the kinase domains of ALKs, hActR-II

(Donaldson *et al.*, 1992; Matzuk and Bradley, 1992), mActR-IIB (Attisano *et al.*, 1992), hTβR-II (Lin *et al.*, 1992) and Daf-1 (Georgi *et al.*, 1990). The percent identity was calculated using Jotun-Hein alignment program (Hein, 1990). (B) Pairwise alignment relationship between the kinase domain of the receptor serine/threonine kinases. The dendrogram was generated using the Jotun-Hein alignment program (Hein, 1990).

Fig. 4. Northern blot analysis of ALK's. A multiple human tissue blot (Clontech) was hybridized with probes for ALK-1 (*Sfi*I-*Bst*XI restriction fragment, corresponding to nucleotides 288 to 670 of ALK-1 cDNA), ALK-2 (*Eco*RI-*Ava*I restriction fragment, including nucleotides 1 to 581 of ALK-2 cDNA), ALK-3 (*Eco*RI-*Pvu*II restriction fragment, corresponding to nucleotides 79 to 824 of ALK-3 cDNA) and ALK-4 (*Sac*I-*Eco*RI restriction fragment, including nucleotides 1178 to 1967 of ALK-4 cDNA). Size markers are indicated on the left.

Fig 1F

```
AAGGAAACGGTTTATTAGGAGGGAGTGGTGGAGCTGGGCCAGGCAGGAAGACGCTGGAATAAGAAACATTTTTGCTCCAGCCCCCATCCCAGTCCCGGGAGGCTGCCGCGCCAGCTGCGC  120
CCAGCGAGCCCCTCCCCGCCTCCAGCCCGGTCCGGGGCCCGCCCCGCACCCCAGCCCGCCGTCCAGCGCTGGCGGTGCAACTGCGGCCCGCCGGTGCAGGGGAGGTGGCCCCGGTCCCCCC  240
AACGTAGCGCCCCCGCCACCCGCAGAGCGGGCCCAGAGGGACCATGACCTTGGGCTCCCCCAGGAAAGGCCTTCTGATGCTGCTGATGGCCTTGGTCACCCAGGGAGACCCTGTGAAGCCG  360
                                                  M  T  L  G  S  P  R  K  G  L  L  M  L  L  M  A  L  V  T  Q  G  D  P  V  K  P   26

TCTCGGGGCCCGCTGGTGACCTGCACGTGTCAGAGCCCACATTGCAAGGGGCCTACCTGCCCGGGGGCCTGGTGCACAGTAGTGCTGGTGCGGGACGAGCGGGAGGCACCCCCAGGAACAT  480
 S  R  G  P  L  V  T  C  T  C  E  S  P  H  C  K  G  P  I  C  R  G  A  W  C  T  V  V  L  V  R  E  E  G  R  H  P  Q  E  H   66

CGGGGCTGCCGCAACTTGCACAGGGAGCTCTGCAGGGGCGCCCCACCGAGTTCGTCAACCACTACTGCTGCGACAGCCACCTCTGCAACCACAACGTGTCCCTGGTGCTGGAGGCCACC  600
 R  G  C  G  N  L  H  R  E  L  C  R  G  R  P  T  E  F  V  N  H  Y  C  C  D  S  H  L  C  N  H  N  V  S  L  V  L  E  A  T  106

CAACCTCCTTCGGAGCAGCCGGGAACAGATGGCCAGCTGGCCCTGATCCTGGGCCCCGTGCTGCCCTTGCTGCCCCTGGGTGCCTGGGCCTGTGGCATGTCCGACGGAGG  720
 Q  P  P  S  E  Q  P  G  T  D  G  Q  L  A  L  F  L  G  P  V  L  A  L  L  A  L  V  A  L  G  V  L  G  L  W  H  Y  R  R  R  148

CAGGAGAAGCAGCGTGGCCTGCACAGCGAGCTGGGAGAGTCCAGTCTCATCCTGAAAGCATCTGAGCAGGGCGACACGATGTTGGGGCACCTCCTGGACAGTGACTGCACCACAGGGAGT  840
 Q  E  K  Q  R  G  L  H  S  E  L  G  E  S  S  L  I  L  K  A  S  E  Q  G  D  T  M  L  G  D  L  L  D  S  D  C  T  T  G  S  188

GGCTCAGGGCTCCCCTTCCTGGTGCAGAGGACAGTGGCACGGCAGGTTGCCTTGGTGGAGTGTGTGGGAAAAGGCCGCTATGGCGAAGTGTGGCGGGGCTTGTGGCACGGTGACAGTGTG  960
 G  S  G  L  P  F  L  V  Q  R  T  V  A  R  Q  Y  A  L  V  E  C  V  G  K  G  R  Y  G  E  V  W  R  G  L  W  H  G  E  S  V  228

GCCGTCAAGATCTTCTCCTCGAGGGATGAACAGTCCTGGTTCCGGGAGACTGAGATCTATAACACAGTATTGCTCAGACACGCACAACATCCTAGGCTTCATCGCCTCAGACATGACCTCC  1080
 A  V  K  I  F  S  S  R  D  E  Q  S  W  F  R  E  T  E  I  Y  N  T  V  L  L  R  H  D  N  I  L  G  F  I  A  S  D  M  T  S  288

CGCAACTCGAGCACGCAGCTGTGGCTCATCACGCACTACCACGAGCACGGCTCCCTCTACGACTTTCTGCAGAGACAGACGCTGGAGCCCCATCTGGCTCTGAGGCTAGCTGTGTCCGCG  1200
 R  N  S  S  T  Q  L  W  L  I  T  H  Y  H  E  H  G  S  L  Y  D  F  L  Q  R  Q  T  L  E  P  H  L  A  L  R  L  A  Y  S  A  308

GCATGCGGCCTGGCCGCACCTGCACGTGGAGATCTTCGGTACACAGGGCAAACCAGCCATTGCCCACCGCGACTTCAAGAGCCGCAATGTGCTGGTCAAGAGCAACCTGCAGTCTTGCATC  1320
 A  C  G  L  A  H  L  H  V  E  I  F  G  T  Q  G  K  P  A  I  A  H  R  D  F  K  S  R  N  V  L  V  K  S  N  L  Q  C  C  I  346

GCCGACCTGGGCCTGGCTGTGATGCACTCACAGGGCAGCGATTACCTGGACATCGGCAACAACCCGACAGTCGGCACCAAGCGGTACATGGCACCCGAGTCCTGGACGAGCAGATCCGC  1440
 A  D  L  G  L  A  V  M  H  S  Q  G  S  D  Y  L  D  I  G  N  N  P  R  V  G  T  K  R  Y  M  A  P  E  V  L  D  E  Q  I  R  386

ACGGACTGCTTTGAGTCCTACAAGTGGACTGACATCTGGGCCTTTGGCCTGGTGCTGTGGGAGATTGCCCGCCGGACCATCGTGAATGGCATCGTGGAGGACTATAGACCACCCTTCTAT  1560
 T  D  C  F  E  S  Y  K  W  T  D  I  W  A  F  G  L  V  L  W  E  I  A  R  R  T  I  V  N  G  I  V  E  D  Y  R  P  P  F  Y  426

GATGTGGTGCCCAATGACCCCAGCTTTGAGCACATGAAGAAGGTGGTGTGTGTGGATCAGCAGACCCCCACCATCCCTAACCGGCTGGCTGCAGACCCGGTCCTCTCAGGCCTACGCTCAG  1680
 D  V  V  P  N  D  P  S  F  E  D  M  K  K  V  V  C  V  D  Q  Q  T  P  T  I  P  N  R  L  A  A  D  P  V  L  S  G  L  A  Q  466

ATGATGCGGGAGTGCTGGTACCCAAACCCCTCTGCCCGACTCACCGCGCTGCGGATCAAGAAGACACTACAAAAAATTAGCAACAGTCCAGAGAAGCCTAAAGTGATTCAATAGCCCAGG  1800
 M  M  R  E  C  W  Y  P  N  P  S  A  R  L  T  A  L  R  I  K  K  T  L  Q  K  I  S  N  S  P  E  K  P  K  V  I  Q  *           503

AGCACCTGATTCCTTTCTGCCTGCAGGGGGCTGGGGGGGTGGGGGGCAGTCGATGGTGCCCTATCTGGGTAGAGGTAGTGTGAGTGTGGTGTGTGCTGGGGATGGGCAGCTGCGCCTGCC  1920
TGCTCGGCCCCCAGCCCACCCAGCCAAAAATACAGCTGGGCTGAAACCTGAAAAAAAAAAAAAA                                                          1984
```

Fig. 1A

Fig. 1B

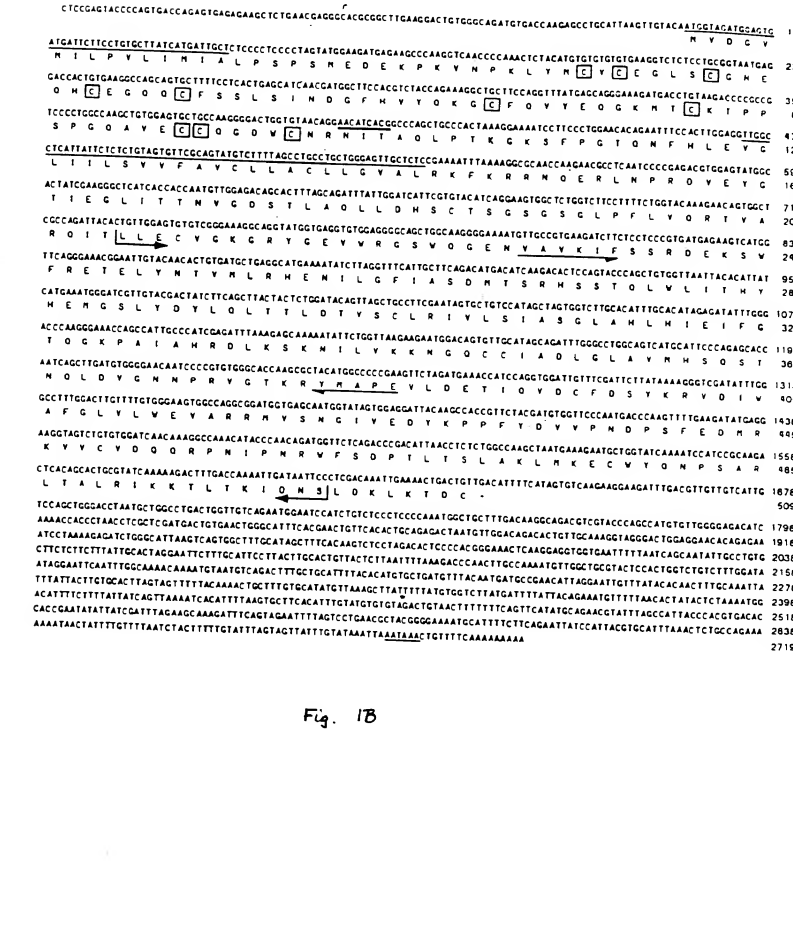
CTCCGAGTACCCCAGTGACCAGAGTGACAGCAAGCTCTGAACGAGGGCACGCGGCTTGAAGGACTGTGGGCAGATGTGACCAACAGCCTGCATTAAGTTGTACAATGGTAGATGGAGTG 118
M V D G V 5
ATGATTCTTCCTGTGCTTATCATGATTGCTCTCCCCTCCCCTAGTATGGAAGATGAGAAGCCCAAGGTCAACCCCAAACTCTACATGTGTGTGTGTGAAGGTCTCTCCTGCGGTAATGAG 238
M I L P V L I M I A L P S P S M E D E K P K V N P K L Y M C Y C C G L S C C N E 45
GACCACTGTGAAGGCCAGCAGTGCTTTTCCTCACTGAGCATCAACGATGGCTTCCACGTCTACCAGAAAGGCTGCTTCCAGGTTTATGAGCAGGGAAAGATGACCTGTAAGACCCCGCCG 358
Q H C E G Q Q C F S S L S I N D G F H V Y Q K G C F Q V Y E Q G K M T C K T P P 85
TCCCCTGGCCAAGCTGTGGAGTGCTGCCAAGGGGACTGGTGTAACAGGAACATCACGGCCCAGCTGCCCACTAAAGGAAAATCCTTCCCTGCAACACAGAATTTCCACTTGGAGGTTGGC 478
S P G Q A Y E C C Q G D W C N R N I T A Q L P T K G K S F P G T Q N F H L E V C 125
CTCATTATTCTCTCTGTAGTGTTCGCAGTATGTCTTTTAGCCTGCCTGCTGGGAGTTGCTCTCCGAAAATTTAAAAGGCGCAACCAAGAACGCCTCAATCCCCGACACGTGGAGTATGGC 598
L I I L S V V F A V C L L A C L L G V A L R K F K R R N Q E R L N P R D V E Y G 165
ACTATCGAAGGGCTCATCACCACCAATGTTGGAGACAGCACTTTAGCAGATTTATTGGATCATTCGTGTACATCAGGAAGTGGCTCTGGTCTTCCTTTTCTGGTACAAAGAACAGTGGCT 718
T I E G L I T T N V G D S T L A Q L L D H S C T S G S G S G L P F L V Q R T V A 205
CGCCAGATTACACTGTTGGAGTCTGTCGGGAAAGGCAGGTATGGTGAGGTGTGCAGGGGCAGCTGGCAAGGGGAAAATGTTGCCGTGAAGATCTTCTCCTCCCGTGATGAGAAGTCATGC 838
R Q I T L L E C V G K G R Y G E V W R G S W Q G E N V A V K I F S S R D E K S W 245
TTCAGGGAAACGGAATTGTACAACACTGTGATGCTGAGGCATGAAAATATCTTAGGTTTCATTGCTTCAGACATGACATCAAGCACTCCAGTACCCAGCTGTGGTTAATTACACATTAT 958
F R E T E L Y N T V M L R H E N I L G F I A S D M T S R H S S T Q L W L I T H Y 285
CATGAAATGGCATCGTTGTACGACTATCTTCAGCTTACTACTCTGGATACAGTTAGCTGCCTTCGAATAGTGCTGTCCATAGCTAGTGGTCTTGCACATTGCACATAGACATATTTGGG 1078
H E M G S L Y D Y L Q L T T L D T V S C L R I V L S I A S G L A H L H I E I F G 325
ACCCAAGGGAAACCAGCCATTGCCCATCGAGATTTAAAGAGCAAAAATATTCTGGTTAAGAAGAATGGACAGTGTTGCATAGCAGATTTGGGCCTGGCAGTCATGCATTCCCAGAGCACC 1198
T Q G K P A I A H R D L K S K N I L V K K N G Q C C I A D L G L A V M H S Q S T 365
AATCAGCTTGATGTGGGGAACAATCCCCGTGTGGGCACCAAGCGCTACATGGCCCCGAAGTTCTAGATGAAACCATCCAGGTGGATTGTTTCGATTCTTATAAAAGGGTCGATATTTGC 1318
N Q L D V G N N P R V G T K R Y M A P E V L D E T I Q V D C F D S Y K R V D I V 405
GCCTTTGCACTTGTTTTGTGGGAAGTGGCCAGGCGGATGCTGAGCAATGGTATAGTGGAGCATTACAAGCCACCGTTCTACGATGTGGTTCCCAATGACCCAAGTTTTGAAGATATGAGG 1438
A F G L V L W E V A R R M V S N G I V E D Y K P P F Y D V V P N D P S F E D M R 445
AAGGTAGTCTGTGTGGATCAACAAAGGCCAAACATACCCAACAGATGGTTCTCAGACCCGACATTAACCTCTCTGGCCAAGCTAATGAAAGAATGCTGGTATCAAAATCCATCCGCAAGA 1558
K V V C V D Q Q R P N I P N R W F S D P T L T S L A K L M K E C W Y Q N P S A R 485
CTCACAGCACTGCGTATCAAAAAGACTTTGACCAAAATTGATAATTCCCTCGACAAATTGAAAACTGACTGTTGACATTTTCATAGTGTCAACAAGGAAGATTTCACGTTGTTGTCATTG 1678
L T A L R I K K T L T K I D N S L D K L K T D C - 509
TCCAGCTGGGACCTAATGCTGCCCTGACTGGTTGTCAGAATGGAATCCATCTGTCTCCCTCCCCAAATGGCTGCTTTGACAAGGCAGACGTCGTACCCAGGCCATGTGTTGGGGAGACATC 1798
AAAACCACCCTAACCTCGCTCGATGACTGTGAACTGGGCATTTCACGAACTGTTCACACTGCAGACACTAATGTTGGACAGACACTGTTGCAAAGGTAGGGACTGGAGGAACACAGACAA 1918
ATCCTAAAAGACATCTGGGCATTAAGTCAGTGGCTTTGCATAGCTTTCACAAGTCTCCTAGACACTCCCCACGGGAAACTCAAGGAGGTGGTGAATTTTTAATCAGCAATATTGCCTGTG 2038
CTTCTCTTCTTTATTGCACTAGGAATTCTTTGCATTCCTTACTTGCACTGTTACTCTTAATTTTAAAGACCCAACTTGCCAAAATGTTGGCTGCGTACTCCACTGGTCTGTCTTTGGATA 2158
ATAGGAATTCAATTTGGCAAAACAAAATGTAATGTCAGACTTTGCTGCATTTTACACATGTGCTGATGTTTACAATGATGCCGAACATTAGGAATTGTTTATACACAACTTTGCAAATTA 2278
TTTATTACTTGTGCACTTAGTAGTTTTACAAAACTGCTTTGTGCATATGTTAAAGCTTATTTTTATGTGCTCTTATGATTTATTACAGAAATGTTTTTAACACTATACTCTAAAATGG 2398
ACATTTTCTTTTATTATCAGTTAAAATCACATTTTAAGTGCTTCACATTTGTATGTGTGTAGACTGTAACTTTTTTTCAGTTCATATGCAGAAGTATTTAGCCATTACCCACGTGACAC 2518
CACCGAATATATTATCGATTAGAAGCAAAGATTTCAGTAGAATTTAGTCCTGAACGCTACGGGGAAAATGCATTTGTCTTCAGAATTATCCATTACGTGCATTTAAACTCTGCCAGAAA 2638
AAAATAACTATTTTGTTTTAATCTACTTTTTGTATTTAGTAGTTATTTGTATAAATTAAATAAACTGTTTTCAAAAAAAAA 2719

Fig. 1C

```
GCTCCGCGCCGAGGGCTGGAGGATGCGTTCCCTGGGTCCGGACTTATGAAAATATGCATCACTTTAATACTCTCTTGGAATTCATGAGATGGAAGCATAGGTCAAAGCTGTTTGGAGAA   120
AATCAGAAGTACAGTTTTATCTAGCCACATCTTGGAGGAGTCGTAACAAAAGCAGTGGGAGTTGAAGTCATTGTCAAGTGCTGCCATCTTTTACAACAAAATCTCACTCAATGATAGTCA   240
TTTAAATTGGTGAAGTAGCAAGACCAATTATTAAAGGTGACAGTACACAGGAAACATTACAATTCAACAATGACTCAGCTATACATTTACATCAGATTATTGGGAGCCTATTTGTTCATC   360
                                                                                                  M  T  G  L  Y  I  Y  I  R  L  L  G  A  Y  L  F  I    17

ATTTCTCGTCTTCAAGCACAGAATCTGGATAGTATGCTTCATGGCACTGGGATGAAATCAGACTCCGACCAGAAAAAGTCAGAAAATGGAGTAACCTTAGCACCAGAGGATACCTTGCCT   480
I  S  R  V  Q  G  N  L  D  S  M  L  H  G  T  G  M  K  S  D  S  Q  Q  K  K  S  E  N  G  V  T  L  A  P  E  Q  T  L  P    57

TTTTTAAAGTGCTATTGCTCAGGGCACTGTCCAGATGATGCTATTAATAACACATGCATAACTAATGGACATTGCTTTGCCATCATAGAAGAAGATGACCAGGGAGAAACCACATTAGCT   600
F  L  K  C  T  C  S  G  H  C  P  Q  Q  A  I  M  N  T  C  I  T  N  G  H  C  F  A  I  I  E  E  D  Q  G  E  T  T  L  A    97

TCAGGGTGTATGAAATATGAAGGATCTGATTTTCAGTGCAAAGATTCTCCAAAAGCCCAGCTACGCCGGACAATAGAATGTTGTCGGACCAATTTATGTAACCAGTATTTGCAACCCACA   720
S  G  C  M  K  Y  E  G  S  D  F  Q  C  K  D  S  P  K  A  Q  L  R  R  T  I  E  C  C  R  T  N  L  C  N  Q  Y  L  Q  P  T   137

CTGCCCCCTGTTGTCATAGGTCCGTTTTTTGATGGCAGCATTCGATGGCTGGTTTTGCTCATTTCTATGGCTGTCTGCATAATTGCTATGATCATCTTCTCCAGCTGCTTTTGTTACAAA   840
L  P  P  V  V  I  G  P  F  F  D  G  S  I  R  W  L  V  L  L  I  S  M  A  V  C  I  I  A  M  I  I  F  S  S  C  F  C  Y  K   177

CATTATTGCAACAGCATCTCAAGCAGACGTCGTTACAATCGTGATTTGCAACAGGATGCAAGCATTTATTCCAGTTGCAGAATCACTAAAAGACCTTATTGACCAGTCACAAAGTTCTGGT   960
H  Y  C  K  S  I  S  S  R  R  R  Y  N  R  D  L  E  Q  D  E  A  F  I  P  V  G  E  S  L  K  D  L  I  Q  Q  S  Q  S  S  G   217

AGTGGGTCTGGACTACCTTTATTGGTTCAGCGAACTATTGCCAAACAGATTCAGATGGTCCGGCAAGTTGGTAAAGGCCGATATGGAGAAGTATGGATGGGCAAATGGCGTGGCGAAAAA  1080
S  G  S  G  L  P  L  L  V  Q  R  T  I  A  K  Q  I  Q  M  V  R  Q  V  G  K  G  R  Y  G  E  Y  W  M  G  K  W  R  G  E  K   257

GTGGCGGTGAAAGTATTCTTTACCACTGAAGAAGCCAGCTGGTTTCGAGAAACAGAAATCTACCAAACTGTGCTAATGCGCCATGAAAACATACTTGGTTTCATAGCGGCAGACATTAAA  1200
V  A  V  K  V  F  F  T  T  E  E  A  S  W  F  R  E  T  E  I  Y  Q  T  V  L  M  R  H  E  N  I  L  G  F  I  A  A  D  I  K   297

GGTACAGGTTCCTGGACTCAGCTCTATTTGATTACTGATTACCATGAAAATGGATCTCTCTATGACTTCCTGAAATGTGCTACACTGGACACCAGAGCCCTGCTTAAATTGGCTTATTCA  1320
G  T  G  S  W  T  Q  L  Y  L  I  T  D  Y  H  E  N  G  S  L  Y  D  F  L  K  C  A  T  L  D  T  R  A  L  L  K  L  A  Y  S   337

GCTGCCTGTGGTCTGTGCCACCTGCACACAGAAATTTATGGCACCCAAGGAAAGCCCGCAATTGCTCATCGACACCTAAAGAGCAAAAACATCCTCATCAAGAAAAATGGGAGTTGCTGC  1440
A  A  C  G  L  C  H  L  H  T  E  I  Y  G  T  Q  G  K  P  A  I  A  H  R  D  L  K  S  K  N  I  L  I  K  K  N  G  S  C  C   377

ATTGCTGACCTGGGCCTTGCTGTTAAATTCAACAGTGACACAAATGAAGTTGATGTGCCCTTGAATACCAGCGTGGGCACCAAACGCTACATGGCTCCCGAAGTGCTGGACGAAAGCCTG  1560
I  A  D  L  G  L  A  V  K  F  N  S  D  T  N  E  V  D  V  P  L  N  T  R  V  G  T  K  R  Y  M  A  P  E  V  L  D  E  S  L   417

AACAAAAACCACTTCCAGCCCTACATCATGGCTGACATCTACAGCTTCGGCCTAATCATTTGGGAGATGGCTCGTCGTTGTATCACACGAGGGATCGTCGAAGAATACCAATTGCCATAT  1680
N  K  N  H  F  Q  P  Y  I  M  A  D  I  Y  S  F  G  L  I  I  W  E  M  A  R  R  C  I  T  G  I  V  E  E  Y  Q  L  P  Y   457

TACAACATGGTACCGAGTGATCCGTCATACGAAGATATGCGTGAGGTTGTGTGTGTCAAACGTTTGCGGCCAATTGTGTCTAATCGGTGGAACAGTCATGAATGTCTACGAGCAGTTTTG  1800
Y  N  M  V  P  S  D  P  S  Y  E  D  M  R  E  V  V  C  V  K  R  L  R  P  I  V  S  N  R  W  N  S  D  E  C  L  R  A  V  L   497

AAGCTAATGTCAGAATGCTGGGCCCACAATCCAGCCTCCAGACTCACAGCATTGAGAATTAAGAAGACGCTTGCCAAGATGGTTGAATCCCAAGATGTAAAAATCTGATCGTTAAACCAT  1920
K  L  M  S  E  C  W  A  H  N  P  A  S  R  L  T  A  L  R  I  K  K  T  L  A  K  M  V  E  S  Q  D  V  K  I  -              532

CGGAGCGAAAACTCTACACTGCAAGAACTGTTTTTACCCATGGCATGGGTGCAATTAGAGTGGAATAAGGATGTTAACTTGGTTCTCACACTCTTTCTTCACTACGTGTCACAGGCTGC  2040
TAATATTAAACCTTTCAGTACTCTTATTAGGATACAAGCTGGGAACTTCTAAACACTTCATTCTTTATATATGGACAGCTTTATTTTAAATCTGGTTTTTGATGGCTTTTTTAAGTGGG  2160
TTTTTATGAACTGCATCAAGACTTCAATCCTGATTAGTGTCTCCAGTCAAGCTCTGGGTACTGAATTGCCTGTTCATAAAACGGTGCTTTCTGTGAAAGCCTTAAGAGATAAATGAGCG  2280
CAGCACAGATGGAGAAATAGACTTTGCCTTTTACCTGAGACATTCAGTTCGTTTGTATTCTACCTTTGTAAAACAGCCTATAGATGATGATGTTTGGCATACTGCTTATTTTATGATA  2400
GTTTGTCCTGTGTCCTTAGTGATGTGTGTGTCTCCATGCACATGCACGCCGGGATTCCTCTGCTGCCATTTGAATTAGAAGAAAATAATTTATATGCATGCACAGGAAGATATTGGTG  2520
GCCGGTGGTTTTGTGCTTTAAAAATGCAATATCTGACCAAGATTCGCCAATCTCATACAAGCCATTTACTTTGCAAGTGAGATAGCTTCCCCACCAGCTTTATTTTTTAACATGAAAGCT  2640
GATGCCAAGGCCAAAAGAAGTTTAAAGCATCTGTAAATTTGCACTGTTTTCCTTCAACCACCATTTTTTTTGTGGTTATTATTTTTGTCACGGAAAGCATCCTCTCCAAAGTTGCAGCTT  2760
CTATTGCCATGAACCATGCTTACAAAGAAAGCACTTCTTATTGAAGTGAATTCCTGCATTTGATAGCAATGTAAGTGCCTATAACCATGTTCTATATTCTTTATTCTCAGTAACTTTTAA  2880
AAGGGAAGTTATTTATATTTTGTGTATAATGTGCTTTATTTGCAAATCACCC                                                                          2932
```

Fig. 1C

Fig. 1D

```
GGCCCGGTGGAGCTGGTAGGCATCATCGCCGGCCCGGTGTTCCTCCTGTTCCTCATCATCATTGTTTTCCTTGTCATTAACTATCATCAGCGTGTCTATCACAACCGCCAGAGACTG  120
 G  P  V  E  L  V  G  I  I  A  G  P  V  F  L  L  F  L  I  I  I  I  V  F  L  Y  I  N  T  H  Q  R  V  Y  H  N  R  Q  R  L   40

GACATGGAAGATCCCTCATGTGAGATGTGTCTCTCCAAAGACAAGACGCTCCAGGATCTTGTCTACGATCTCTCCACCTCAGGGTCTGGCTCAGGGTTACCCCTCTTTGTCCAGCGCACA  240
 D  M  E  D  P  S  C  E  M  C  L  S  K  D  K  T  L  Q  D  L  V  Y  D  L  S  T  S  G  S  G  S  G  L  P  L  F  V  Q  R  T   80

GTGGCCCGAACCATCGTTTTACAAGAGATTATTGGCAAGCGTCGGTTTGGGCAAGTATGGCGGGGCCGCTGGAGGGGTGGTGATGTGGCTGTGAAAATATTCTCTTCTCGTGAAGAACGG  360
 V  A  R  T  I  V  L  Q  E  I  I  G  K  G  R  F  G  E  V  W  R  G  R  W  R  G  G  D  V  A  V  K  I  F  S  S  R  E  E  R  120

TCTTGGTTCAGGGAAGCAGAGATATACCAGACGGTCATGCTGCGCCATGAAAACATCCTTGGATTTATTGCTGCTGACAATAAAGATAATGGCACCTGGACACAGCTGTGGCTTGTTTCT  480
 S  W  F  R  E  A  E  I  Y  Q  T  V  M  L  R  H  E  N  I  L  G  F  I  A  A  D  N  K  D  N  G  T  W  T  Q  L  W  L  V  S  160

GACTATCATGAGCACGGGTCCCTGTTTGATTATCTGAACCGGTACACAGTGACAATTGAGGGGATGATTAAGCTGGCCTTGTCTGCTGCTAGTGGGCTGGCACACCTGCACATGGAGATC  600
 D  T  H  E  H  G  S  L  F  D  T  L  N  R  Y  T  V  T  I  E  G  M  I  K  L  A  L  S  A  A  S  G  L  A  H  L  H  M  E  I  200

GTGGGCACCCAAGGGAAGCCTGGAATTGCTCATCGAGACTTAAAGTCAAAGAACATTCTGGTGAAGAAAAATGGCATGTGTGCCATAGCAGACCTGGGCCTGGCTGTCCGTCATGATGCA  720
 V  G  T  Q  G  K  P  G  I  A  H  R  D  L  K  S  K  N  I  L  V  K  K  N  G  M  C  A  I  A  D  L  G  L  A  Y  R  H  D  A  240

GTCACTGACACCATTGACATTGCCCCGAATCAGAGGGTGGGGACCAAACGATACATGGCCCCTGAAGTACTTGATGAAACCATTAATATGAAACACTTTGACTCCTTTAAATGTGCTGAT  840
 V  T  D  T  I  D  I  A  P  N  Q  R  V  G  T  K  R  Y  M  A  P  E  V  L  D  E  T  I  N  M  K  H  F  D  S  F  K  C  A  D  280

ATTTATGCCCTCGGGCTTGTATATTGGGAGATTGCTCGAAGATGCAATTCTGGAGGAGTCCATGAAGAATATCAGCTGCCATATTACGACTTAGTGCCCTCTGACCCTTCCATTGAGGAA  960
 I  T  A  L  G  L  V  T  V  E  I  A  R  R  C  N  S  G  G  V  H  E  E  T  Q  L  P  T  T  D  L  V  P  S  P  S  I  E  E  320

ATGCGAAAGGTTGTATGTGATCAGAAGCTGCGTCCCAACATCCCCAACTGGTGGGCAGTTATGAGGCACTGCGGGTGATGGGGAAGATGATGCGAGAGTGTTGGTATGCCAACGGCGCA 1080
 M  R  K  V  Y  C  D  Q  K  L  R  P  N  I  P  N  W  W  G  S  T  E  A  L  R  V  M  G  K  M  M  R  E  C  W  Y  A  N  G  A  360

GCCCGCCTGACGGCCCTGCGCATCAAGAAGACCCTCTCCCAGCTCAGCGTGCAGGAAGACGTGAAGATCTAACTGCTCCCTCTCTCCACACGGAGCTCCTGGCAGCGAGAACTACGCACA 1200
 A  R  L  T  A  L  R  I  K  K  T  L  S  Q  L  S  V  Q  E  D  V  K  I  *                                                   383

GCTGCCGCGTTCAGCGTACGATGGAGGCCTACCTCTCGTTTCTGCCCAGCCCTCTGTGGCCAGGAGCCCTGGCCCGCAAGAGGGACAGAGCCCGGGAGAGACTCGCTCACTCCCATGTTG 1320
GGTTTGAGACAGACACCTTTTCTATTTACCTCCTAATGGCATGGAGACTCTGAGAGCGAATTGTGTGGAGAACTCAGTGCCACACCTCGAACTGCTTGTAGTGGGAAGTCCCGCGAAACC 1440
CGGTGCATCTGGCACGTGGCCAGGAGCCATGACAGGGGCGCTTGGGAGGGGCCGGAGGAACCGAGGTGTTGCCAGTGCTAAGCTGCCCTGAGGGTTTCCTTCGGGGACCAGCCCACAGCA 1560
CACCAAGGTGGCCCGGAAGAACCAGAAGTGCAGCCCCTCTCACAGGCAGCTCTGAGCCGCGCTTTCCCCTCCTCCCTGGGATGGACGCTGCCGGGAGACTGCCAGTGGAGACGAAATCTG 1680
CCGCTTTGTCTGTCCAGCCGTGTGTGCATGTGCCGAGGTGCGTCCCCCGTTGTGCCTGGTTCGTCCATGCCCTTACACGTGCGTGTGACTGTGTGTGTGTCTGTAGGTGCGCACTTA 1800
CCTGCTTGAGCTTTCTGTGCATGTGCAGCTCGGGGGTGTGGTCGTCATGCTGTCCGTGCTTGCTGGTGCCTCTTTTCAGTAGTGAGCAGCATCTAGTTTCCCTGGTGCCCTTCCCTGGAG 1920
GTCTCTCCCTCCCCCAGAGCCCCTCATGCCACAGTGGTACTCTGTGT                                                                        1967
```

Fig. 1D

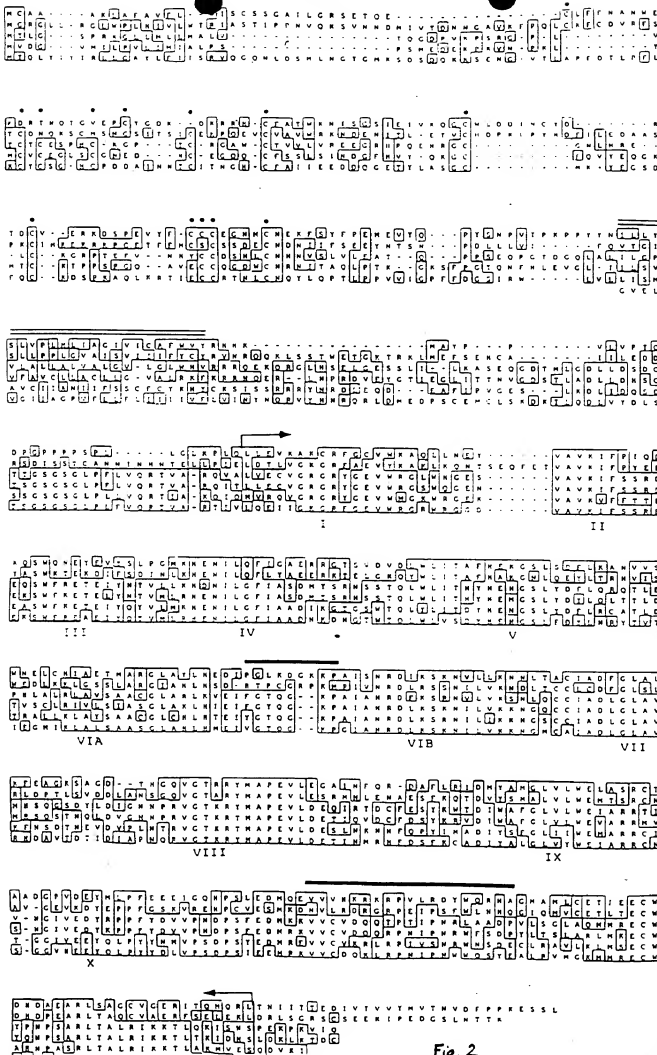

Fig. 2

Fig 3A

Percent Identity in kinase domains

ALK-2	ALK-3	ALK-4	ALK-5	ActR-II	ActR-IIB	TβR-II	daf-1	
79	60	61	63	40	40	37	39	ALK-1
	63	64	65	41	39	37	39	ALK-2
		63	65	41	38	37	39	ALK-3
			90	41	40	39	42	ALK-4
				42	40	41	43	ALK-5
					78	48	35	ActR-II
						47	32	ActR-IIB
							34	TβR-II

Fig. 3A

Fig 3B

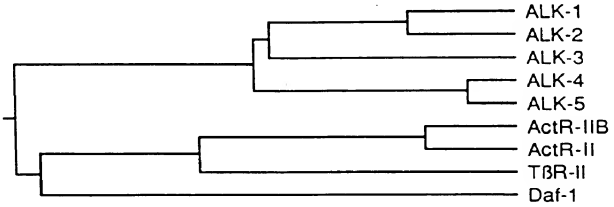

Fig. 3B

Fig 4

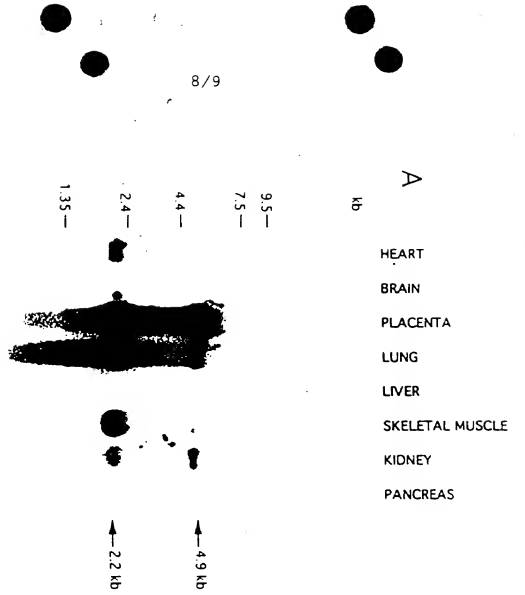

Fig. 4A

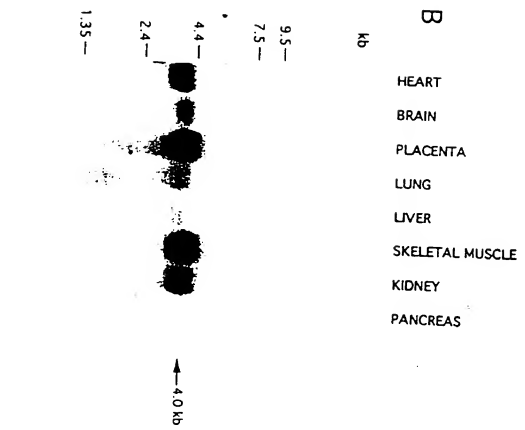

Fig. 4B

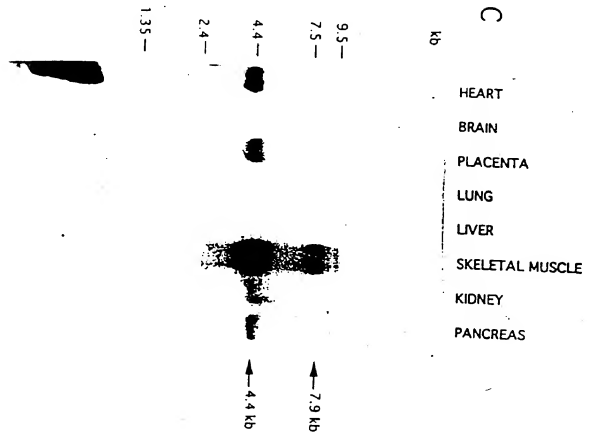
C
kb
9.5—
7.5—
4.4—
2.4—
1.35—
HEART
BRAIN
PLACENTA
LUNG
LIVER
SKELETAL MUSCLE
KIDNEY
PANCREAS
4.4 kb
7.9 kb

Fig. 4C

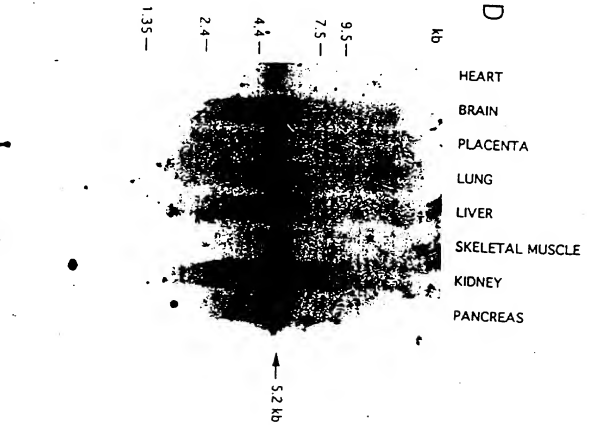
D
kb
9.5—
7.5—
4.4—
2.4—
1.35—
HEART
BRAIN
PLACENTA
LUNG
LIVER
SKELETAL MUSCLE
KIDNEY
PANCREAS
5.2 kb

Fig. 4D